

# IMAGE RETRIEVAL SYSTEM

Publication number: JP2003187217

Publication date: 2003-07-04

Inventor: TAKEUCHI KENJI

Applicant: NEF KK

Classification:

- international: G06T1/00; G06F17/30; G06T7/00; G06T1/00;  
G06F17/30; G06T7/00; (IPC1-7): G06T1/00; G06F17/30;  
G06T7/00

- European:

Application number: JP20010386843 20011220

Priority number(s): JP20010386843 20011220

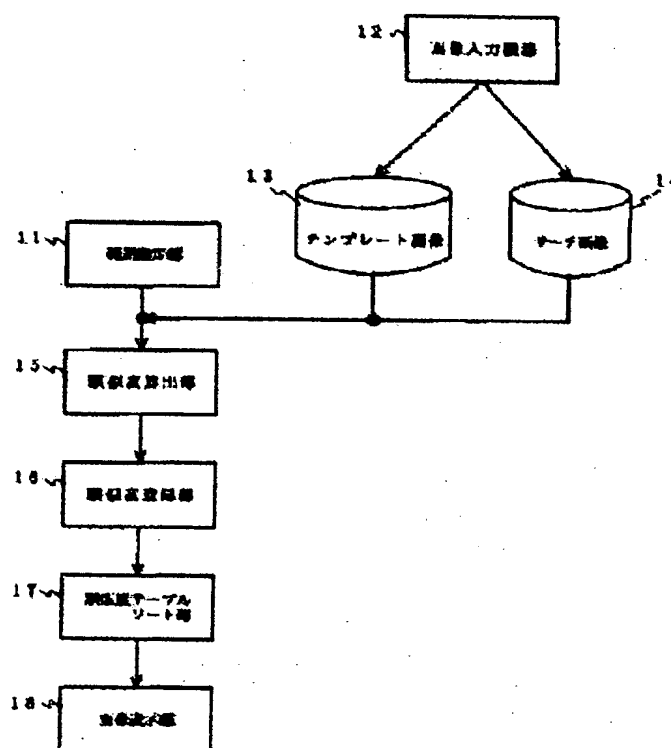
Report a data error here

## Abstract of JP2003187217

**PROBLEM TO BE SOLVED:** To provide an image retrieval system by which an image fetched by a facsimile or a scanner, etc., can be retrieved without using a keyword or ruled lines.

**SOLUTION:** A range specification part 11 specifies a range of the image when similarity between a template image 13 and search images 14 is calculated. A degree of similarity calculation part 15 calculates degrees of similarity between the template image 13 and the search images 14 in a specified image range. A degree of similarity registration part 16 registers the degree of similarity (a numerical value) calculated by the similarity calculation part 15 in a degree of similarity table. A degree of similarity table sorting part 17 sorts (rearranges) the degrees of similarity of the search images 14 to an optional template image 13 in a descending order. An image display part 18 displays corresponding search images 14 in an order from the one with the highest degree of similarity.

COPYRIGHT: (C)2003,JPO



Data supplied from the esp@cenet database - Worldwide

(19) 日本国特許庁 (J P)

## (12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2003-187217

(P2003-187217A)

(43) 公開日 平成15年7月4日 (2003.7.4)

(51) Int.Cl. <sup>7</sup>	識別記号	F I	フォーマット*(参考)
G 0 6 T 1/00	2 0 0	C 0 6 T 1/00	2 0 0 E 5 B 0 5 0
G 0 6 F 17/30	1 7 0	C 0 6 F 17/30	1 7 0 B 5 B 0 7 5
	3 1 0		3 1 0 Z 5 L 0 9 6
	3 5 0		3 5 0 C
G 0 6 T 7/00	3 0 0	C 0 6 T 7/00	3 0 0 E
審査請求 未請求 請求項の数3 O L (全 9 頁)			

(21) 出願番号 特願2001-386843(P2001-386843)

(22) 出願日 平成13年12月20日 (2001. 12. 20)

(71) 出願人 000232047

株式会社エヌ・イー・エフ

東京都港区芝五丁目7番1号

(72) 発明者 竹内 健司

東京都港区芝浦三丁目18番21号 日本電気

エンジニアリング株式会社内

(74) 代理人 100088812

弁理士 ▲柳▼川 信

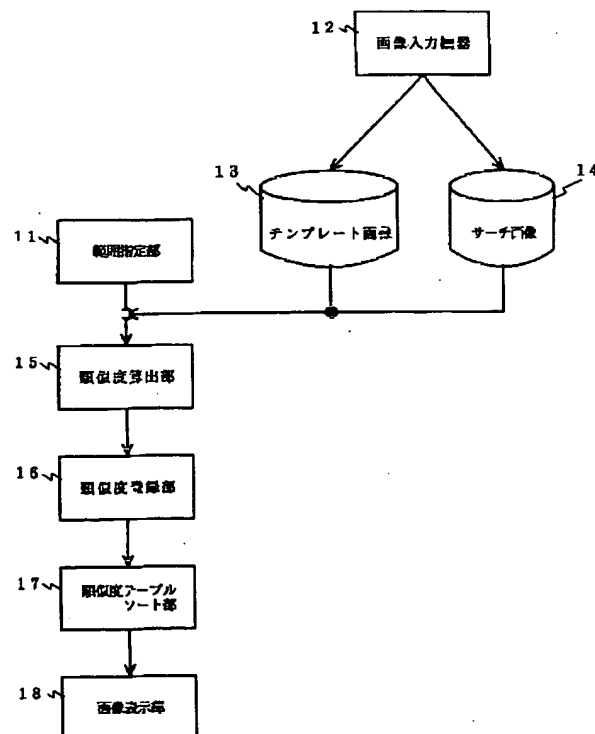
最終頁に続く

## (54) 【発明の名称】 画像検索システム

## (57) 【要約】

【課題】 ファクシミリやスキャナ等で取り込まれた画像を、キーワードや野線を用いることなく検索可能な画像検索システムを提供する。

【解決手段】 範囲指定部11はテンプレート画像13とサーチ画像14との類似度を算出する際の画像の範囲を指定する。類似度算出部15は指定された画像範囲においてテンプレート画像13とサーチ画像14との類似度を算出する。類似度登録部16は類似度算出部15で算出された類似度(数値)を類似度テーブルに登録する。類似度テーブルソート部17は任意のテンプレート画像13に対するサーチ画像14の類似度を降順にソート(並べ替え)を行う。画像表示部18は類似度が高い順に対応するサーチ画像14を表示する。



## 【特許請求の範囲】

【請求項1】 画像の雛形であるテンプレート画像と外部から入力されるサーチ画像との類似度を算出する際に対象とする画像範囲を指定する範囲指定手段と、前記範囲指定手段で指定された画像範囲における前記テンプレート画像と前記サーチ画像との類似度を算出する類似度算出手段と、

前記類似度算出手段で算出された類似度を類似度テーブルに登録する類似度登録手段と、

任意のテンプレート画像に対するサーチ画像毎に前記類似度テーブルに登録された類似度を降順に並べ替える類似度テーブルソート手段と、

前記類似度テーブルソート手段で並べ替えられた類似度が高い順に対応するサーチ画像を表示する画像表示手段とを有することを特徴とする画像検索システム。

【請求項2】 画像の雛形であるテンプレート画像から認識するための情報が含まれている局所範囲を検出する局所範囲検出手段と、

前記局所範囲検出手段で検出された局所範囲における前記テンプレート画像と外部から入力されるサーチ画像との類似度を算出する類似度算出手段と、

前記類似度算出手段で算出された類似度を類似度テーブルに登録する類似度登録手段と、

任意のテンプレート画像に対するサーチ画像毎に前記類似度テーブルに登録された類似度を降順に並べ替える類似度テーブルソート手段と、

前記類似度テーブルソート手段で並べ替えられた類似度が高い順に対応するサーチ画像を表示する画像表示手段とを有することを特徴とする画像検索システム。

【請求項3】 印刷及び記入のいずれかの帳票の画像を前記テンプレート画像及び前記サーチ画像に電子データとして保存する画像入力機器を含むことを特徴とする請求項1または請求項2記載の画像検索システム。

## 【発明の詳細な説明】

## 【0001】

【発明の属する技術分野】本発明は画像検索システムに関し、特に画像識別技術を用いた画像検索システムに関する。

## 【0002】

【従来の技術】従来、画像識別システムや画像処理システムとしては、特開平5-217019号公報（第1の公報）に記載されているように、ビジネスフォームの罫線に着目し、線の数、長い線の数、交点座標等を特徴として抽出するシステムがある。

【0003】上記のシステムでは、テンプレートの特徴を予めフォーム辞書として登録し、検索対象となるサーチ画像に関しても同様の特徴抽出を行い、その特徴抽出の結果と辞書との比較によって目的の画像を検索している。

【0004】また、従来の画像検索システムにおいて

は、画像にキーワード等を付加して登録しておき、検索時にはキーワードの組み合わせによって目的の画像を検索するシステムもある。

【0005】さらに、従来の画像検索システムとしては、特開平8-279019号公報（第2の公報）に記載されているように、使用するフォームシートを予めフォーム生成・登録部で生成しておき、ファクシミリ等で送られてきたノイズを含むような帳票であっても、そのフォーマットを検出可能とするシステムもある。

## 【0006】

【発明が解決しようとする課題】しかしながら、上述した従来のシステムでは、上記の第1の公報記載のシステムの場合、検索する画像内に罫線があることを前提としているため、検索する画像に罫線のない画像（ビジネスフォーム）に対して適用することができない。

【0007】また、検索時にキーワードの組み合わせを用いる場合には、画像に付加するキーワードの命名に一定の基準がなく、また画像の内容を全て網羅するキーワードまたは文書表現が不可能に近いため、画像に付加したキーワード検索においてキーワードの付け方や検索時のキーワードの指定によって目的とする画像が大量に検索されてしまったり、目的の画像を検索することができないことがある。

【0008】さらに、上記の第2の公報記載のシステムの場合、連結黒画素の最小外接矩形を作成し、矩形情報を作成しているため、検出対象の帳票が罫線で囲まれていると、フォームの特徴抽出を正常に行うことができない。この場合、矩形情報作成処理を行うと、罫線で囲まれた帳票のほぼ全面が特徴量となってしまう、罫線で囲まれた内部の個別文字位置矩形を検出することができない。

【0009】さらにまた、上記の第2の公報記載のシステムの場合、文字列の位置、大きさ等を予め登録し、その登録内容を用いて該当する帳票を検出しているため、使用するフォームシートデザインが予めフォーム生成・登録部にてシステム設計者が実施する必要があり、自動化されていない。

【0010】そこで、本発明の目的は上記の問題点を解消し、ファクシミリやスキャナ等で取り込まれた画像を、キーワードや罫線を用いることなく検索することができる画像検索システムを提供することにある。

【0011】また、本発明の他の目的は、検索者が予め検索画像の特徴となる範囲を指定しなくても自動的にテンプレート画像から検索する範囲を検出して登録することができる画像検索システムを提供することにある。

## 【0012】

【課題を解決するための手段】本発明による画像検索システムは、画像の雛形であるテンプレート画像と外部から入力されるサーチ画像との類似度を算出する際に対象とする画像範囲を指定する範囲指定手段と、前記範囲指

定手段で指定された画像範囲における前記テンプレート画像と前記サーチ画像との類似度を算出する類似度算出手段と、前記類似度算出手段で算出された類似度を類似度テーブルに登録する類似度登録手段と、任意のテンプレート画像に対するサーチ画像毎に前記類似度テーブルに登録された類似度を降順に並べ替える類似度テーブルソート手段と、前記類似度テーブルソート手段で並べ替えられた類似度が高い順に対応するサーチ画像を表示する画像表示手段とを備えている。

【0013】本発明によるたの画像検索システムは、画像の雛形であるテンプレート画像から認識するための情報が含まれている局所範囲を検出する局所範囲検出手段と、前記局所範囲検出手段で検出された局所範囲における前記テンプレート画像と外部から入力されるサーチ画像との類似度を算出する類似度算出手段と、前記類似度算出手段で算出された類似度を類似度テーブルに登録する類似度登録手段と、任意のテンプレート画像に対するサーチ画像毎に前記類似度テーブルに登録された類似度を降順に並べ替える類似度テーブルソート手段と、前記類似度テーブルソート手段で並べ替えられた類似度が高い順に対応するサーチ画像を表示する画像表示手段とを備えている。

【0014】すなわち、本発明の画像検索システムは、テンプレート画像とサーチ画像との類似度を算出する際の画像の範囲を指定する範囲指定部と、印刷または記入された帳票の画像をテンプレート画像またはサーチ画像の電子データとして保存する画像入力機器と、検索者が求めている画像の雛形であるテンプレート画像とスキャナやファクシミリ等の画像入力機器から入力される多種多様な画像であるサーチ画像とに対して範囲指定部で指定された画像範囲だけをテンプレート画像とサーチ画像との類似度の算出対象とする類似度算出部と、算出された類似度（数値）を類似度テーブルに登録する類似度登録部と、任意のテンプレート画像に対するサーチ画像毎に登録された類似度に対して降順にソート（並べ替え）を行う類似度テーブルソート部と、類似度が高い順に対応するサーチ画像を表示する画像表示部とを備えることを特徴としている。

【0015】また、本発明の画像検索システムでは、範囲指定部の代わりに、検索者が検索画像の特徴となる範囲を指定しなくても自動的にテンプレート画像から局所範囲を検出する局所範囲検出部を備えることを特徴としている。

【0016】上記のように、目的とする画像と類似した画像を登録（画像入力）し、その特徴的な位置を指定して画像を検索することによって、多種多様な画像にキーワード等を付加せずに格納しない場合や文書内に野線を含まない画像においても目的の画像検索が可能となることである。

【0017】また、テンプレート画像とサーチ画像との

類似度計算を全範囲で行わずに局所的に行うことによって、画像検索処理時間を短縮することが可能となる。

【0018】さらに、局所範囲検出部を実装した場合、テンプレート画像を登録（入力）するだけでテンプレート画像内の下地（白）以外の範囲、つまり予め印字されている文字や野線等の範囲を特定することができ、その範囲の類似度を自動的に算出して類似度テーブルに登録することによって、実際に検索者が検索を行う場合、類似度を改めて算出する必要が少なくなる。

【0019】

【発明の実施の形態】次に、本発明の実施例について図面を参照して説明する。図1は本発明の一実施例による画像検索システムの構成を示すブロック図である。図1において、本発明の一実施例による画像検索システムは範囲指定部11と、画像入力機器12と、類似度算出部15と、類似度登録部16と、類似度テーブルソート部17と、画像表示部18とから構成されている。

【0020】範囲指定部11はテンプレート画像13とサーチ画像14との類似度を算出する際の画像の範囲を指定する。画像入力機器12は印刷または記入された帳票の画像をテンプレート画像13またはサーチ画像14に電子データとして保存する。この場合、テンプレート画像13は検索者が求めている画像の雛形であり、サーチ画像14はスキャナやファクシミリ等の画像入力機器12から入力される多種多様な画像である。

【0021】類似度算出部15はテンプレート画像13とサーチ画像14との類似度を算出する。その際、類似度算出部15は範囲指定部11で指定された画像範囲だけを算出対象とする。類似度登録部16は類似度算出部15で算出された類似度（数値）を類似度テーブル（図示せず）に登録する。類似度テーブルソート部17は任意のテンプレート画像13に対するサーチ画像14の類似度を降順にソート（並べ替え）を行う。画像表示部18は類似度が高い順に対応するサーチ画像14を表示する。

【0022】図2は図1のテンプレート画像13とサーチ画像14との具体例を示す図である。図2において、テンプレート画像2-1はFax送付表であり、帳票の種類2-5、ロゴマーク2-4、送付先2-6、送付元2-7等が予め印刷されている。

【0023】例えば、検索者が自分宛に届いたFaxの画像を検索する場合、帳票の種類を示す局所範囲（1）2-2と、送付先を示す局所範囲（2）2-3とを指定する。指定された局所範囲内（矩形内部）の画像と類似した画像を含むものをサーチ画像2-14の中から検索する。

【0024】テンプレート画像2-8もFax送付表である。この帳票の場合、帳票種類2-10を示す範囲は上部にあり、ロゴマーク2-12は下部にあり、全体を囲む様に飾り野線2-13がある。

【0025】検索者がこの帳票と類似した画像を検索する場合を考える。同じ種類の帳票ではあるが、中身に記述される内容2-11はそれぞれ異なっているため、除外範囲(1)2-9の矩形外の画像が類似した画像をサーチ画像2-14の中から検索する。

【0026】図3は本発明の一実施例による画像検索システムの動作を示すフローチャートである。これら図1～図3を参照して本発明の一実施例による画像検索システムの動作について説明する。

【0027】最初に検索者は、自分が検索したい画像が既にテンプレート画像13内にあればそれを選択し、テンプレート画像13内になければ画像入力機器12を使用し画像入力する(図3ステップS1)。

【0028】次に、検索者は検索する画像に特徴的な局所範囲2-2、2-3や除外範囲2-9を指定する(図3ステップS2)。この指定されたテンプレートや範囲指定が以前検索された実績があるか否かを判断する(図3ステップS3)。検索実績がある場合にはステップS7に遷移し、検索実績がない場合にはサーチ画像14を1件入力する(図3ステップS4)。

【0029】類似度算出部15は入力したサーチ画像14とテンプレート画像13との類似度を指定範囲内だけで算出する(図3ステップS5)。類似度登録部16は類似度算出部15で算出された類似度を該当サーチ画像の類似度テーブル(図6参照)に登録する(図3ステップS6)。

【0030】類似度テーブルソート部17は算出した類似度が予め指定した閾値より大か否か(テンプレート画像に似ているか否か)を判断する(図3ステップS7)。類似度テーブルソート部17は算出した類似度が閾値以下の場合にステップS10に遷移し、閾値より大の場合に類似度テーブルを類似度が大きい順番にソートする(図3ステップS8)。画像表示部18は類似度が大きい順に画面上にサーチ画像を表示する(図3ステップS9)。

【0031】続いて、類似度算出を行っていないサーチ画像が残っているか否かを判断し(図3ステップS10)。残っている場合にはステップS3から処理を繰り返す。サーチ画像が残っていない場合には検索処理を終了する。

【0032】図4は局所範囲が指定されたテンプレート画像のデータ内容を示す図である。図4において、局所範囲が指定されたテンプレート画像は登録テンプレート画像4-1と、検索者が指定した局所範囲(1)4-2及び局所範囲(2)4-3とから構成されている。

【0033】局所範囲は矩形の左上及び右下の座標データとする。局所範囲(1)は(xmin1, ymin1, xmax1, ymax1)となり、局所範囲(2)は(xmin2, ymin2, xmax2, ymax2)となる(図4の4-4参照)。類似度を算出する場

合には、テンプレート画像とサーチ画像の局所範囲

(1)、局所範囲(2)とでそれぞれ正規化相関係数等の類似度計算を行う。

【0034】図5は除外範囲が指定されたテンプレート画像のデータ内容を示す図である。図5において、除外範囲が指定されたテンプレート画像は登録テンプレート画像5-1と、検索者が指定した除外範囲5-2とから構成されている。除外範囲は矩形の左上及び右下の座標データとする。除外範囲は(xmin3, ymin3, xmax3, ymax3)となる(図5の5-3参照)。類似度を算出する場合には、テンプレート画像とサーチ画像の除外範囲とは正規化相関係数等の類似度計算を行わない。除外範囲外(矩形の外側)のみ類似度計算の対象とする。

【0035】図6は類似度テーブル例を示す図である。図6において、類似度テーブルの各レコードは、サーチ画像ファイル名称、テンプレート画像名称、範囲指定のフィールドから構成されている。1レコードがサーチ画像1件に対応する。

【0036】また、図6に示す類似度テーブル内の数字は、図2のテンプレート2-1において局所範囲(1)及び局所範囲(2)が指定された時の類似度と、図2のテンプレート2-8において除外範囲が指定された時の類似度とが登録されていることを示している。

【0037】この例では、類似度は0以上、1以下の数値として表しており、値が大きい(1に近い)程類似した画像であることを示している。検索実績の有無判定(図3ステップS3)、類似度登録(図3ステップS6)、降順ソート(図3ステップS8)の処理はこの類似度テーブルを利用する。

【0038】図7は本発明の他の実施例による画像検索システムの構成を示すブロック図である。図7において、本発明の他の実施例による画像検索システムは範囲指定部11の代わりに局所範囲検出部19を配設した以外は図1に示す本発明の一実施例による画像検索システムの構成と同様であり、同一構成要素には同一符号を付してある。また、同一構成要素の動作は本発明の一実施例と同様である。

【0039】局所範囲検出部19は検索者が検索画像の特徴となる範囲を指定しなくても、自動的にテンプレート画像から認識するための情報が含まれている局所範囲を検出する。

【0040】図8は図7の局所範囲検出部19を用いてテンプレート画像2-1から局所範囲を検出する動作について説明するための図である。この図8を参照して局所範囲検出部19がテンプレート画像2-1から局所範囲を検出する動作について説明する。

【0041】まず、局所範囲検出部19は横方向及び縦方向のヒストグラムを検出する。この操作によって、テンプレート画像2-1のどこに情報(白い下地以外のデ

ータ)が含まれているかを特定することができる。

【0042】すなわち、テンプレート画像2-1においてはロゴが印字されている位置8-1、帳票種類を示す印字8-2、送付先8-3、送付元8-4それぞれを囲む矩形位置が検出され、これらを局所範囲とする。この局所範囲の検出処理以降の処理については、上述した本発明の一実施例と同様である。

【0043】このような構成とすることによって、検索者がいない場合においても、テンプレート画像を登録するだけで類似度計算を行い、類似度テーブルを作成することができる。つまり、実際に検索者が検索を行う前に登録したテンプレートに関してはサーチ画像が分類されていることと同様になる。より詳細な検索時、例えば自分宛にきたFaxを検索する場合等、予め分類されたサーチ画像の中から更に局所範囲を指定することによって高速で確実な検索が可能となる。つまり、検索者が予め検索画像の特徴となる範囲を指定しなくても自動的にテンプレート画像から検索する範囲を検出して登録することができる。

【0044】このように、本発明においては、目的とする画像と類似した画像を登録(画像入力)し、その特徴的な位置を指定して画像を検索することによって、多種多様な画像にキーワード等を付加せずに格納しない場合や文書内に罫線を含まない画像においても、目的の画像検索を行うことができる。

【0045】また、本発明では、テンプレート画像とサーチ画像との類似度計算を全範囲で行わずに局所的に行うことによって、画像に対する検索処理時間を短縮することができる。

【0046】さらに、本発明では、局所範囲検出部19を実装した場合、テンプレート画像を登録(入力)するだけでテンプレート画像内の下地(白)以外の範囲、つまり予め印字されている文字や罫線等の範囲を特定することができる、その範囲の類似度を自動的に算出して類似度テーブルに登録することによって、実際に検索者が検索を行う場合、類似度を改めて算出する必要が少なくなる。

【0047】

【発明の効果】以上説明したように本発明の画像検索システムは、画像の雛形であるテンプレート画像と外部から入力されるサーチ画像との類似度を算出する際に対象とする画像範囲を指定し、その指定された画像範囲におけるテンプレート画像とサーチ画像との類似度を算出し、算出された類似度を類似度テーブルに登録し、任意のテンプレート画像に対するサーチ画像毎に類似度テーブルに登録された類似度を降順に並べ替え、この並べ替えられた類似度が高い順に対応するサーチ画像を表示することによって、ファクシミリやスキャナ等で取り込まれた画像を、キーワードや罫線を用いることなく検索することができるという効果が得られる。

【0048】また、本発明の画像検索システムは、画像の雛形であるテンプレート画像から認識するための情報が含まれている局所範囲を検出し、その検出された局所範囲におけるテンプレート画像と外部から入力されるサーチ画像との類似度を算出し、算出された類似度を類似度テーブルに登録し、任意のテンプレート画像に対するサーチ画像毎に類似度テーブルに登録された類似度を降順に並べ替え、この並べ替えられた類似度が高い順に対応するサーチ画像を表示することによって、検索者が予め検索画像の特徴となる範囲を指定しなくても自動的にテンプレート画像から検索する範囲を検出して登録することができるという効果が得られる。

【図面の簡単な説明】

【図1】本発明の一実施例による画像検索システムの構成を示すブロック図である。

【図2】図1のテンプレート画像とサーチ画像との具体例を示す図である。

【図3】本発明の一実施例による画像検索システムの動作を示すフローチャートである。

【図4】局所範囲が指定されたテンプレート画像のデータ内容を示す図である。

【図5】除外範囲が指定されたテンプレート画像のデータ内容を示す図である。

【図6】類似度テーブル例を示す図である。

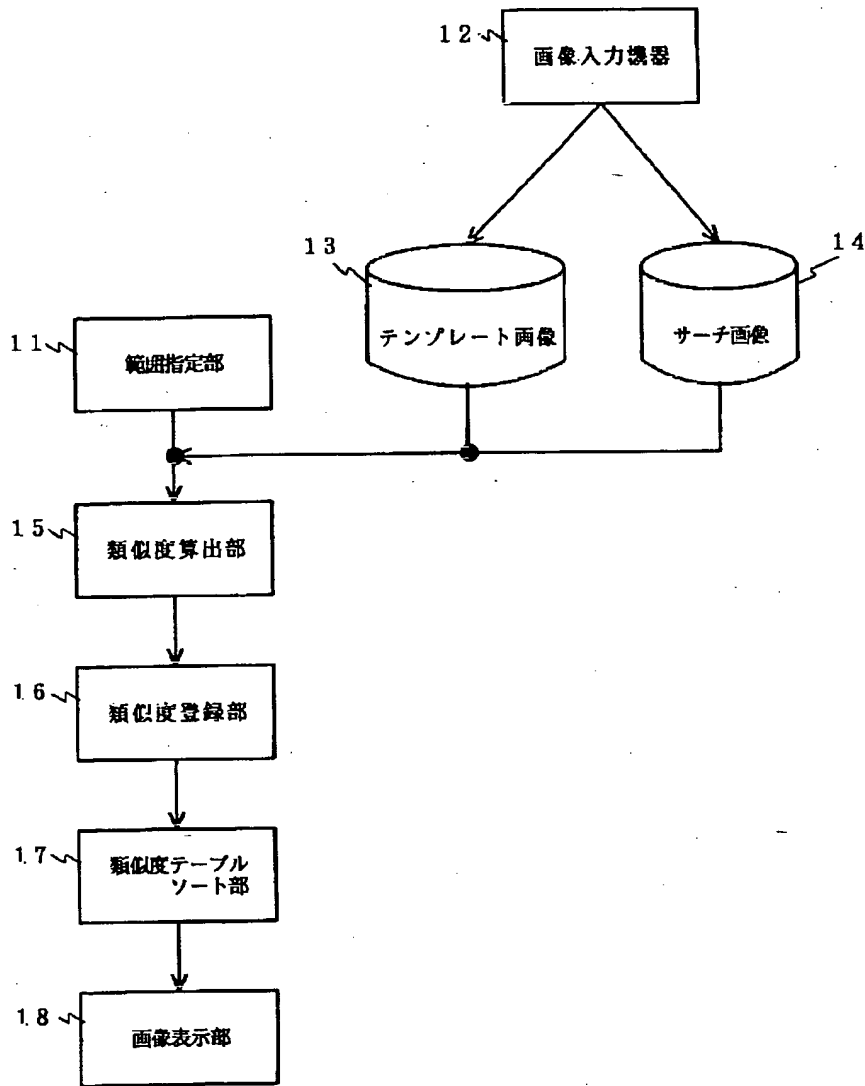
【図7】本発明の他の実施例による画像検索システムの構成を示すブロック図である。

【図8】図7の局所範囲検出部を用いてテンプレート画像から局所範囲を検出する動作について説明するための図である。

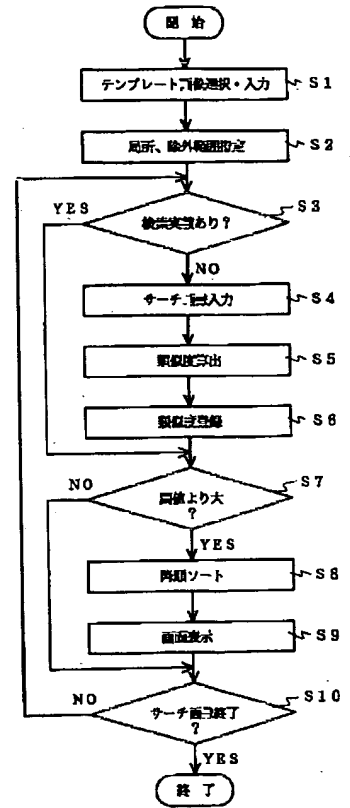
【符号の説明】

- 2-1, 2-8, 13 テンプレート画像
- 2-2 局所範囲(1)
- 2-3 局所範囲(2)
- 2-4, 2-12 ロゴマーク
- 2-5, 2-10 帳票の種類
- 2-6, 8-3 送付先
- 2-7, 8-4 送付元
- 2-9 除外範囲(1)
- 2-13 飾り罫線
- 2-14, 14 サーチ画像
- 8-1 ロゴが印字されている位置
- 8-2 帳票種類を示す印字
- 11 範囲指定部
- 12 画像入力機器
- 15 類似度算出部
- 16 類似度登録部
- 17 類似度テーブルソート部
- 18 画像表示部
- 19 局所範囲検出部

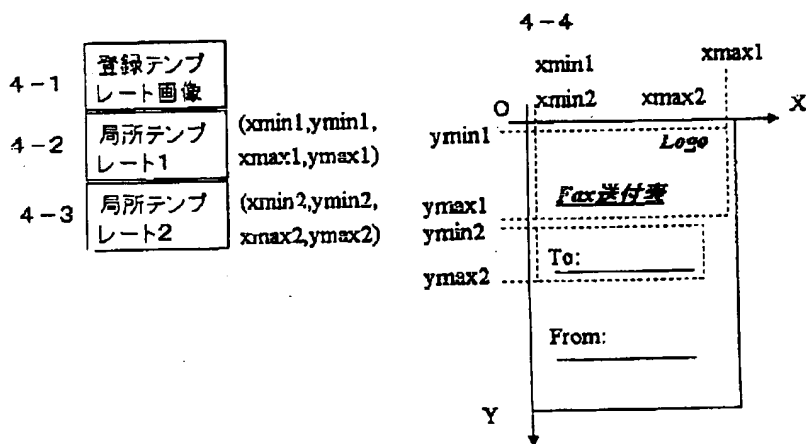
【図1】



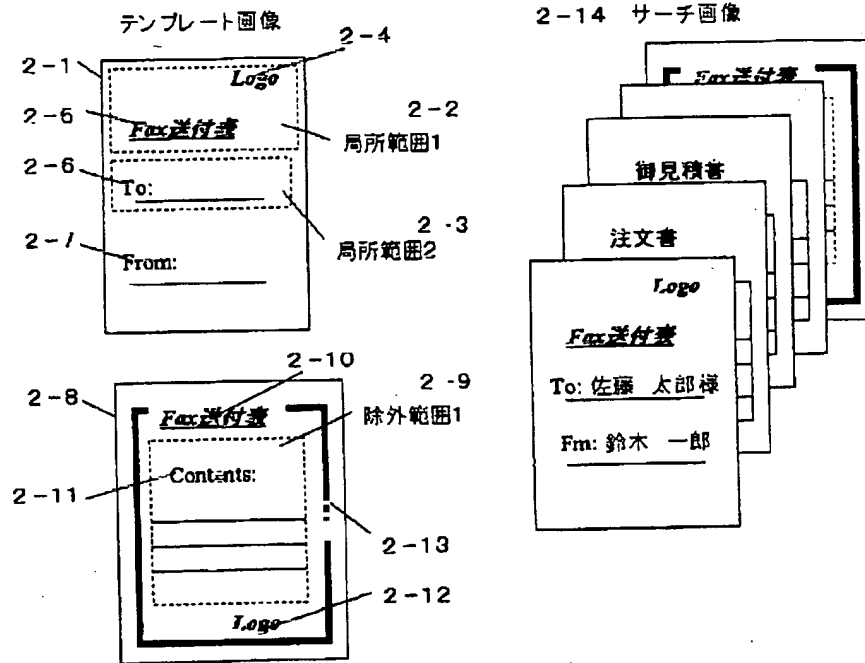
【図3】



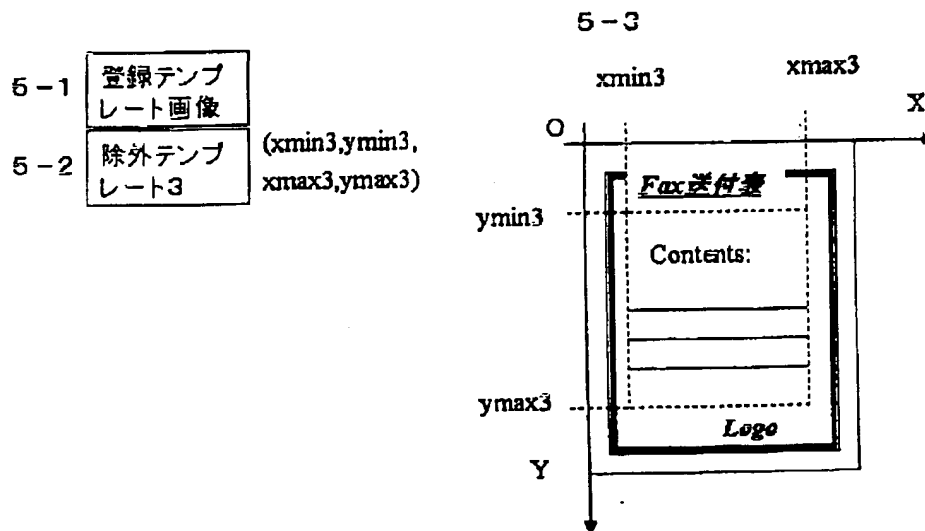
【図4】



【図2】



【図5】

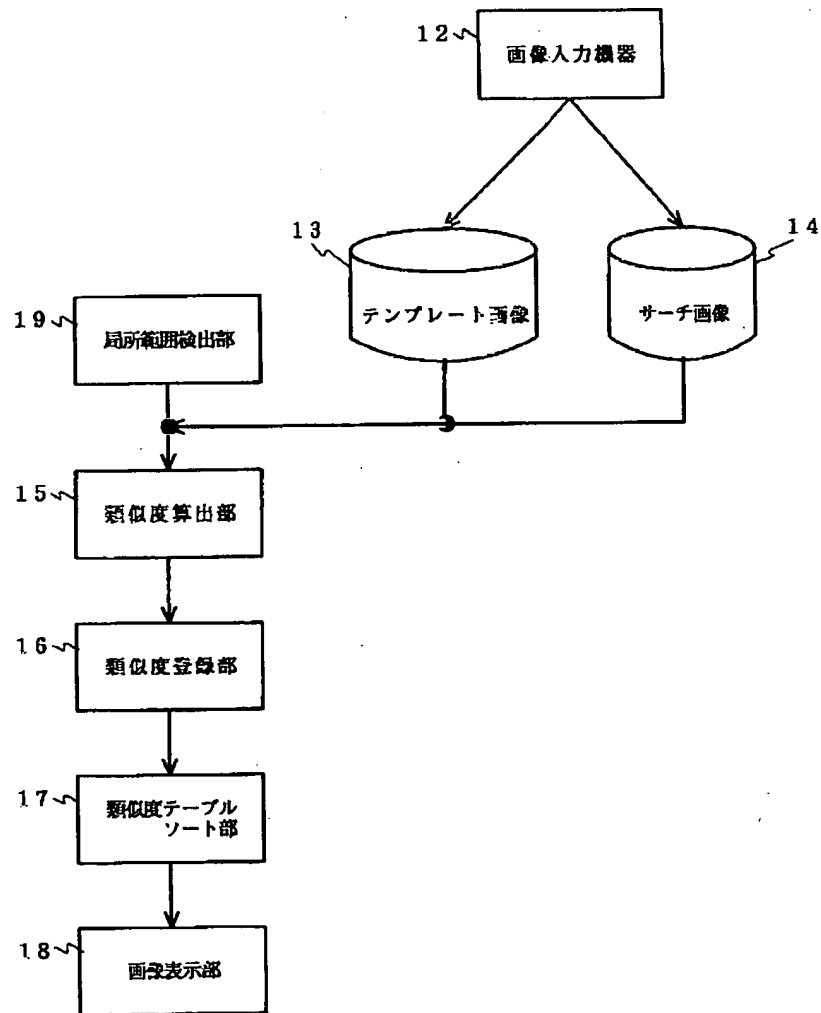




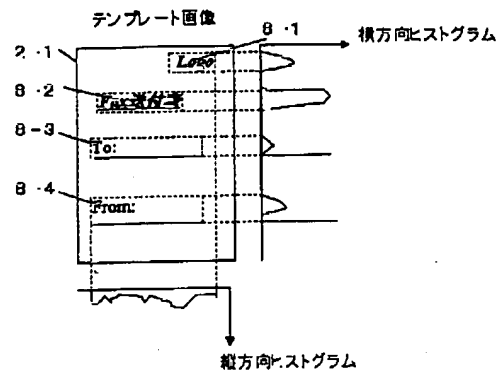
【図6】

サーチ画像 ファイル名	テンプレート2-1		テンプレート2-8
	局所画素#1	局所画素#2	除外範囲
Image#1	0.7	0.4	0.1
Image#2	0.7	0.7	0.2
Image#3	0.2	0.1	0.8
⋮	⋮	⋮	⋮
Image#n	0.3	0.2	0.1

【図7】



【図8】



フロントページの続き

Fターム(参考) 5B050 AA10 BA10 BA16 CA07 DA06  
 EA18 FA02 FA09 FA12 FA19  
 GA08  
 5B075 ND06 PP28 QM05  
 5L096 AA03 AA07 BA07 BA17 CA18  
 CA24 EA37 FA34 HA08 JA03  
 KA11 KA15